WHITHAM, CURTIS & CHRISTOFFERSON, P.C.
INTELLECTUAL PROPERTY LAW
Providing Global Intellectual Property Strategies & Solutions

Michael E. Whitham
Marshall M. Curtis
Clyde R. Christofferson
Robert N. Cook
*C. Lamont Whitham**
*Mary E. Goulet**
**of counsel*

Registered Patent Agents:
Ruth E. Tyler-Cross, Ph. D
Olga V. Merkulova
Daniel A. Steinberg

Law Clerk:
Philana S. Handler

FAX COVER PAGE

DATE: April 11, 2005 **TIME:** 10:19 pm

TO: Examiner Abul K. Azad (703-872-9306)
Art Unit 2654

FROM: Clyde Christofferson

SUBJECT: 09/605,709

8 PAGES, INCLUDING COVER PAGE

REPLY BRIEF UNDER 37 C.F.R. §1.193(b)

Return fax #: 703-787-7557; if all pages were not correctly received please call 703-787-9400

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re patent application of

Frederick J. Damerau and David E. Johnson

Serial No. 09/605,709 Group Art Unit 2654

Filed June 27, 2000 Examiner Abul K. Azad

For AUTOMATED SET UP OF WEB-BASED NATURAL LANGUAGE INTERFACE

Mail Stop Appeal Brief – Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

CERTIFICATE OF TRANSMISSION UNDER 37 CFR §1.8

Date of Transmission: April 11, 2005

I hereby certify that the foregoing REPLY BRIEF OF APPELLANTS UNDER 37 C.F.R. §1.193(b) in connection with the patent application of Frederick J. Damerau and David E. Johnson for AUTOMATED SET UP OF WEB-BASED NATURAL LANGUAGE INTERFACE is being transmitted by facsimile to the USPTO at 703-872-9306, under 37 C.F.R. §1.8 and in accordance with the notice "Patent Customers Advised to FAX Communications to USPTO" dated 11/2/2001 and posted on the USPTO web site http://www.uspto.gov/september11/faxnotice.htm, and as modified by the instruction "Centralized Delivery and Facsimile Transmission Requirements for Patent Application Related Correspondence" issued 10/1/2003.

Clyde R Christoffersen

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re patent application of

Frederick J. Damerau and David E. Johnson

Serial No. 09/605,709 Group Art Unit 2654

Filed June 27, 2000 Examiner Abul K. Azad

For AUTOMATED SET UP OF WEB-BASED NATURAL LANGUAGE INTERFACE

Mail Stop Appeal Brief – Patents
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

REPLY BRIEF OF APPELLANTS UNDER 37 C.F.R. §1.193(b)

Sir:

The Examiner mailed an Answer on February 10, 2005 responsive to Appelleant's Supplemental Appeal Brief. Please charge Deposit account 50-0510 of International Business Machines Corporation (IBM-Yorktown) in the amount of $500.00 (37 C.F.R. 1.17(c)) to cover the fee for filing this reply brief.

FURTHER ARGUMENT

The Examiner argues in his Answer that Sarukkai teaches sparse n-grams. For the reasons that follow, the applicant respectfully disagrees with this conclusion.

It is important to understand what it is that Sarukkai discloses. Sarukkai discloses an improved speech interface to the Internet. The Internet is not like other speech recognition domains, such as when a person reads text or speaks spontaneously. The Internet presents a different problem from other speech

recognition domains because of the huge vocabulary of the Internet. As with other speech recognition domains, the user is trained on a vocabulary, thereby enabling the speech recognition system to use the spoken representation to search for and find the appropriate textual representation. This is done within a model space, and n-grams are commonly used as the language model for this space.

But for a user trying to use voice commands to search the Internet, the size of the Internet vocabulary means an increased likelihood that the user will speak using words which are "out-of-vocabulary" in the terminology of voice recognition. The problem being addressed by Sarukkai is finding the textual representations (i.e. text words, from a vocabulary) that match the spoken representation. This is done by searching, **but not the kind of searching that is being done by the user who is using voice commands to navigate the Internet.** The "speech recognition search" specified by Sarukkai (col. 2, lines 64-65; col. 3, line 2; col. 7, line 21) is trying to find the corresponding text for spoken words.

However, as Sarukkai states explicitly, this "search" for the optimal word sequence W (i.e. textual representation) corresponding to a given acoustic observation X combines acoustic scores (for acoustic observation X) and language scores (for the possible word sequence W) in an evaluation function (col. 4, lines 55-65). This is a highly mathematical definition of a "search" – but is well understood in the art of speech recognition. In non-mathematical terms, what is going on here is the search for a best fit out of possible text word sequences to a given voice sequence. The process starts from the observed voice sequence, which is given an acoustic score, and then evaluates possible text word sequences (each having a language score) in combination with this acoustic score using an evaluation function, ending up with a best fit. Sarukkai's contribution to this voice recognition evaluation is to bias the

evaluation function towards "the set of words that are present in the web page (or that specify links) currently being viewed" (col. 5, lines 7-14). The results of this bias is to improve the ability of the voice recognition system to correctly identify the web page desired from 45% to 75%, as shown in Table 2 of Sarukkai.

The distinction between the present invention and Sarukkai may be better understood by considering an example that appears in both disclosures, and by considering the significance of n-grams, and in particular "sparse n-grams", in dealing with this example. As will be clear from what follows, Sarukkai takes a completely different approach to the example. The example is discussed in the background section of Sarukkai (col. 3, lines 30-36), in connection with the limitations of certain prior art approaches. Suppose that the phrase "THE CURRENT STOCK QUOTES" is the only link in the page, and the user speaks the words "THE CURRENT STOCK PRICE QUOTES." If the word "PRICE" is missing from the voice recognition vocabulary, the prior art method (RGDAG) would fail to make the connection.

Now consider how Sarukkai characterizes and solves this problem. Sarukkai notes that a proper method of handling these "out of vocabulary" (OOV) words is not available and the problem is even more prominent in the context of web surfing (col. 2, lines 1-16). Significantly, Sarukkai notes in the background section that "it is necessary to get the in-vocabulary words correct in the presence of OOV words" (col. 2, lines 14-16). This is significant because this is precisely the effect of the biasing technique used by Sarukkai, which boosts the probabilities of **individual** words belonging to the web triggered word set (col. 10, lines 26-31; emphasis supplied). There is nothing indicated for **the sequence** of words in the web triggered word set: it is the probabilities of **individual** words that are boosted in the evaluation function.

Thus, this method of building on a language model (col. 10, lines 23-24) easily caters to integration with n-grams (col. 10, line 35).

Now consider the present invention. As stated in the Supplemental Appeal Brief (at page 9), the term "sparse n-gram" refers to **the sequence** of words allowing for gaps between words making up the n-gram. The gaps are limited by establishing a distance d which is the maximum separation between the first and last words of the n-gram. Thus, the above example "THE CURRENT STOCK PRICE QUOTES" may be represented as the 4-gram "THE CURRENT STOCK QUOTES" with a d value of 4 instead of 3.

The Examiner argues that "sparse n-gram" has no meaning because it is claimed using the language "wherein the n-grams may be sparse or non-sparse." The Examiner argues that the phrase "sparse or non-sparse" is inclusive and therefore any kind of n-grams can be read on the claimed limitations. This is a slight of hand argument. It is clear from the disclosure of the present invention that the inventors have defined the term "sparse n-gram" with its distance d to allow for gaps. The n-grams referred to in Sarukkai are the conventional n-grams where $d = n - 1$, i.e. there are no gaps. The Examiner's slight of hand argument essentially tries to read the claim language out of the claim. This cannot rightly be done. If sparse n-grams have a gap, i.e. $d > n - 1$, then the language "sparse or non-sparse" means $d \geq n - 1$. In essence, by claiming "sparse or non-sparse" the clear meaning of the inventors is to use an expanded notion of n-grams, including "sparse n-grams" which allow for gaps. While it is possible to rephrase the language to say "wherein the n-grams include sparse n-grams" the meaning would not change.

Further, the Examiner cites the teaching from Sarukkai regarding "a set of words selectively extracted from the Web page source" (col. 7, lines 65-66) as

referring to a sparse n-gram, but this is incorrect. The cited reference is not to an n-gram but rather to the above described "set of words that are present in the web page" and used to bias **the probabilities of individual words** in the evaluation function. Consequently, as Sarukkai makes clear, this "set of words" is not an n-gram (sparse or otherwise) but rather is a collection of individual words for a boosting approach that builds upon the n-gram language model.

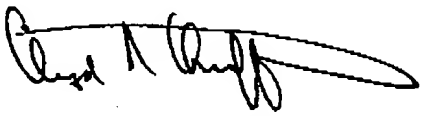
The Examiner attempts to obtain further mileage from the "set of words" used by Sarukkai for biasing as described above. The Examiner argues that this "set of words" amounts to a "natural language interface" as claimed by the invention. It is manifest that the Examiner's suggestion in this regard cannot stand. The "set of words" is a collection of words used individually for biasing the voice recognition evaluation function. Although these words are taken from a web page, which the user can observe visually, any function they have on the web page because of their sequence or arrangement is lost when they are collected into a mere "set of words" for biasing purposes. They have no function in Sarukkai relating to an interface, much less a natural language interface. Indeed, Sarukkai provides a "speech interface" (col. 1, line 16) to information on the Internet, whereas the present invention is concerned not with speech mediated by a voice recognition system but with natural language as that term is used by those skilled in the computer arts, namely, to distinguish languages used by humans for general-purpose communication from constructs such as computer programming languages or the languages used in the study of formal logic.

CONCLUSION

The issue for resolution in this appeal is whether claims 1 to 6 are anticipated by U.S. Patent No. 5,819,220 to Sarukkai under the objective standards of 35 U.S.C. §102(b). The Supplemental Appeal Brief, further supplemented by the foregoing arguments, demonstrates that Sarukkai does not anticipate each and every element of the claimed invention.

In view of the foregoing, it is respectfully submitted that the final rejection of claims 1-6 is in error. Accordingly, reversal of the final rejection is respectfully requested.

Respectfully submitted,



Clyde R Christofferson
Reg. 34,138
Attorney for Applicants

Whitham, Curtis, & Christofferson, P.C.
11491 Sunset Hills Road, Suite 340
Reston, VA 20190
703-787-9400
703-787-7557 (fax)